

AD-A129 848

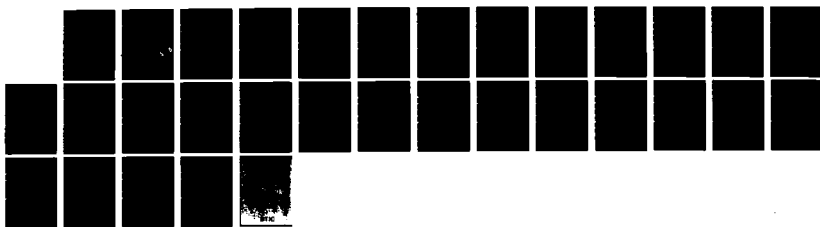
MEASUREMENT OF TASK PERFORMANCE ON THE BASIS OF THE
TIME/EVENT MATRIX: AN. (U) MILTON S HERSHEY MEDICAL
CENTER HERSHEY PA DEPT OF BEHAVIORAL.. S STREUFERT
MAR 83 TR-12-ONR N00014-80-C-0581

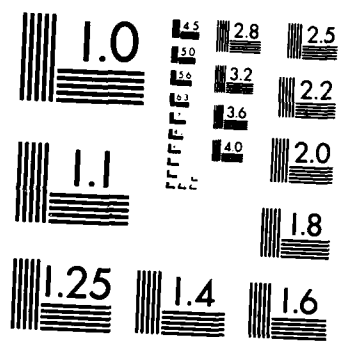
1/1

UNCLASSIFIED

F/G 5/8

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963 A

unclassified

①

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report ONR No. 12	2. GOVT ACCESSION NO. AD-A129848	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Measurement of Task Performance on the Basis of the Time/Event Matrix: An Extension of Methods.		5. TYPE OF REPORT & PERIOD COVERED
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Siegfried Streufert, Ph.D.		8. CONTRACT OR GRANT NUMBER(s) N00014-80-C-0581
9. PERFORMING ORGANIZATION NAME AND ADDRESS Pennsylvania State University, College of Medicine, Dept. of Behavioral Science, Milton S. Hershey Medical Center, Hershey PA 17033		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 170 - 909
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research, Code 442 Quincy Street, Arlington VA 22217		12. REPORT DATE March 1983
		13. NUMBER OF PAGES 28
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution is unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) complexity, decision making, decision matrix, differentiation, information, integration, measurement, planning quality of decisions, quantity of decisions, respondent behavior, retaliatory behavior, simulation, speed, strategy, stress, time/event matrix, task performance, uncertainty.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Previously reported measures of complex task performance are supplemented with a number of new measures. Formulas for calculating all measures now utilized are presented. Data collection on the basis of the time/event matrix is explained. Information about the print-out numbers of the various measures for some recently developed complex man-machine simulations is provided.		

DTIC
SELECTED
JUN 29 1983
H

ADA 129848

DTIC FILE COPY



Accession No.	
NTIS GTR	
DTIC TAB	
Unannounced	
Justification	
By	
Distribution/	
Availability Code	
Dist	Avail and/or Special
A	

Measurement of Task Performance on the Basis of
the Time-Event Matrix: An Extension of Methods

Siegfried Streufert

The Pennsylvania State University, College of Medicine

In a previous technical Report (ONR No. 3, 1981), Streufert and Streufert presented, among other material, seven formulas designed to measure components of complex decision making. These formulas were based on two decades of development, representing orthogonal measures of decision making processes (in contrast to decision content). Validity information for these measures has been reported by a number of authors.

With the advent of a complex experimental micro-computer-assisted simulation to assess task performance, it appeared useful to expand our current measures and to include a number of additional scoring techniques. Such process measurement, as it has been developed to date, will be discussed in this report.

The "decision matrix" discussed in the previous technical report is, in effect, a time/event matrix which can be quite useful to assess a number of performance processes (in addition to, or instead of, complex decision making). Further, the matrix may be extended from the current two dimensions to three or more dimensions, allowing for the inclusion of additional variables. The basic time variable in the two dimensional matrix does, of course, remain in all cases. However, while the decision variable may continue to be utilized, a third variable might, for example, be communications, etc. The measurement formulas that will be presented in this report are based on a two-dimensional matrix, but can easily be amended to allow for three or more-dimensional measurement. In the following pages, the time/event matrix and

fifteen measures based on it will be discussed in considerable detail. The measures include the original seven presented in Technical Report ONR No. 3, although one of these (Spread across Decision Categories) has been slightly modified on the basis of recent data collection experience. Print-out score number of computer based time/event matrix scoring for two recently developed micro-computer assisted simulations* are provided to assist those who wish to use that research vehicle. Details about the simulation itself will be presented in a subsequent technical report.

PURPOSE OF THE TIME/EVENT MATRIX

The task of an individual or a task-oriented group operating in the world outside of the laboratory is rarely limited to deciding on a single event within a limited context. For example, most decision makers in applied settings must respond to an ongoing series of inputs from their environment. The resulting output is usually a sequence of actions determined in part by some plan and in another part by the necessities of dealing with current events. The output may consist of primarily "respondent" actions or it may reflect some degree of "strategy", i.e. decisions which are interrelated with each other and occur in a planned sequence to achieve some kind of goal. Whether or not individual or group actions do reflect pure respondent behavior, whether they reflect some kind of strategy (and the level and/or characteristics of that strategy) may be of considerable importance for the outcome of the task effort. The majority of previous researchers have not focused on measuring or describing such differences. To alleviate that problem, Streufert and associates have developed a time/event matrix to help researchers or observers to identify the different kinds of actions and their frequencies, as they occur in naturally complex task settings. Reliability and validity of these measurement techniques have been established in previous efforts.

The following pages will describe how time/event matrices are constructed on the basis of individual or group performance. Subsequently a number of formulas describing performance measures will be provided and explained. The formulas based on the matrix (to be presented in the last section of this paper) have been shown to reflect generally orthogonal

*developed at Pennsylvania State University under this contract with ONR and at Science Applications, Inc., McLean, VA under a contract with ARI.

measurement of performance style and performance outcome. Additional measures may be developed on the basis of the time/event matrix where needed.

THE MATRIX

Performance quality, particularly in complex tasks, is determined by at least two components of the individual or group effort: (1) appropriate knowledge about what responses are potentially correct or incorrect (where possible) and (2) the ability to develop a plan, to respond at the right time and with optimal combination of responses, i.e. the use of strategy. The time/event matrix was designed to measure the latter of the two components. In many cases, the first component, i.e., appropriate content knowledge and understanding of the task situation can be assumed, as long as sufficient training and experience is available. However, persons with training and experience can differ widely on the second characteristic.

Time/event matrices can be used to measure a variety of task performance activities, depending on the interests and orientations of the researcher or observer. This paper cannot cover all of the purposes for which the matrix can or has been employed. For greater ease of communication, we will focus on decision-making matrices as an example for all matrix possibilities. It should be remembered, however, that most other performance areas, aside from decision-making, could have been selected equally well.

The time/event matrix technique was primarily developed to measure the interrelationships among actions over time and the effects of information flow which precedes those actions. As suggested above, the matrix is not sensitive to the content of actions (e.g., decisions) and will not distinguish between "correct" and "incorrect" actions. If quality of content is of concern, additional measures (beyond those discussed here) will have to be selected to obtain complete information about overall performance levels. However (as suggested earlier), where personnel is well trained and/or experienced, that will frequently not be necessary.

Details about the construction of time/event matrices will be discussed below. At this point it is merely important to be aware that these matrices capture all data about incoming information, about decisions and other actions based on that information, about interrelationships among information and decisions as well as interrelationships among decisions (e.g. strategy). These matrices may be used to collect data on a number of measures which reflect how task oriented individuals or groups process information and how that information processing determines or affects observed performance. Measures based on the time/event matrices may be considered "intermediate" assessments of performance quality. They provide a necessary vehicle for estimating and defining important action antecedents of performance (criterion based) quality, particularly in complex tasks.

Tasks and their requirements differ. The same strategy is not necessarily useful in all task settings (aside from differences in knowledge content). As a result, the measures derived from the time/event matrix should be carefully validated against each general performance task. Many tasks will, of course, produce quite similar patterns of "optimal" measurement levels to criterion. As a matter of fact, a number of validations have shown that specific score levels for the measures that will be presented later tend to be quite robust across a number of tasks and a number of performance environments. Where adequate training and/or sufficient experience is likely to result in few (if any) content errors in performance, predictions of quality task performance made on the basis of matrix measurement is, consequently of substantial value. In addition, the matrix may be expanded into 3(or n-) dimensional space, permitting measurement and prediction of additional relevant variables as they impact on, or interact with, task performance and the measures of task performance and that will be discussed below.

ESTABLISHING THE MATRIX

The two dimensions of a matrix are time and action (here decision) types. Each will be discussed in turn:

(1) Time.

Time in the matrix is plotted horizontally. There are no particular restrictions on the gradations to be used (no matter whether time proceeds normally or is - as in some simulations - expanded or condensed), except that events which occur sequentially and independently of each other must appear on different time points. The time dimension moves from the left to the right. The units of the scale used are not of significance, except that decision-making sequences which are to be compared must contain the same scale units (since the formulas should calculate comparable values).

(2) Decision Types.

Decision-making tasks and settings differ. Consequently the types of decisions employed differ as well. For example, executives dealing with the potential purchase of another corporation may be concerned with such action areas as establishing the value of the other company, determining potential duplication of effort, etc. On the other hand, military decision makers may be concerned with troop movements, air support decisions and so forth. In other words, groupings of decisions (decision types) must be established separately for each general group of decision-making situations. Selection of decision-making types is best done by experts in the field. The types selected should be inclusive; where possible of approximately equal breadth, and conceptually meaningful and consistent. The types should clearly differ from each other in activity, method, meaning, etc. Decision types should provide the potential for use by the decision maker. While some decision makers would likely use one group of decisions, others may use a different group, of course with considerable overlap.

While there is no restriction on the number of potential decision types that might be represented in a time/event matrix, decision types should be selected so that decision makers utilize, on the average, somewhere between ten and twenty different types of decisions¹ in any time sequence that lasts for several hours. Note, however, that these sugges-

¹Since decision makers would rarely employ all available decision types, the potential for considerably more than 10-20 decision types (as in more general terms action categories) may be provided.

tions are ideal requirements and do not supercede the practical characteristics of decision-making situations. For example, if a decision situation requires only one kind of decision, one cannot "manufacture" other decision types by hook or crook. In effect, the use of the decision matrix in such simple situations would not be of any value. For example, if all actions reflected troop movements, then splitting decisions by the unit moved may not be of value.

(3) Decision Points.

Once time is plotted horizontally and decision types (as selected, for example, by an expert panel) are plotted vertically, each decision made by an individual or a group of decision makers (as desired by the researcher or observer) can be represented by a point placed vertically underneath the time when that decision was made (or announced, or transmitted, again depending on the intent of the researcher or observer) and horizontally next to the decision type represented. All decisions can be so placed in the matrix. Decisions made at the same point in time may be connected with vertical lines. Decisions representing the same decision type may be connected with horizontal lines.

(4) Information Input.

In the matrix, as used so far, information input is only considered as it relates to decision output (this limitation was chosen for convenience and is not necessary). Any unit of input which leads to an output is marked (e.g. by a *) under its appropriate (input) time and in front of (on the same decision-type line as) any decision made as a consequence of that input. The input is placed in advance of each output which it produced, i.e. it may occur on more than one horizontal (decision-type) line. The distance on the horizontal between the input * and the decision point reflects the time elapsed between receipt of information and relevant response.²

(5) Diagonals

As stated above, we are interested in the relationships among decisions as they reflect, for example, the development of plans or

²See the section on Integration Time Weight (below) for a discussion of time measurement.

strategies. Consequently we wish to know whether a decision made at one time is related (leads) to a decision at a later time. Where a decision of one decision type is made to make a later decision of another type possible, the two decisions are connected across time with a diagonal line with an arrow-head pointing forward toward the later decision.³ If two decisions show an isolated relationship to each other, a single arrow is drawn. If, on the other hand, the decision maker(s) decides to engage in decision types A and B at time one to allow for action C in the future, and wants to accomplish C to allow D to occur even later, and if all these decisions are actually made in time, a longer chain of diagonal connections is established:

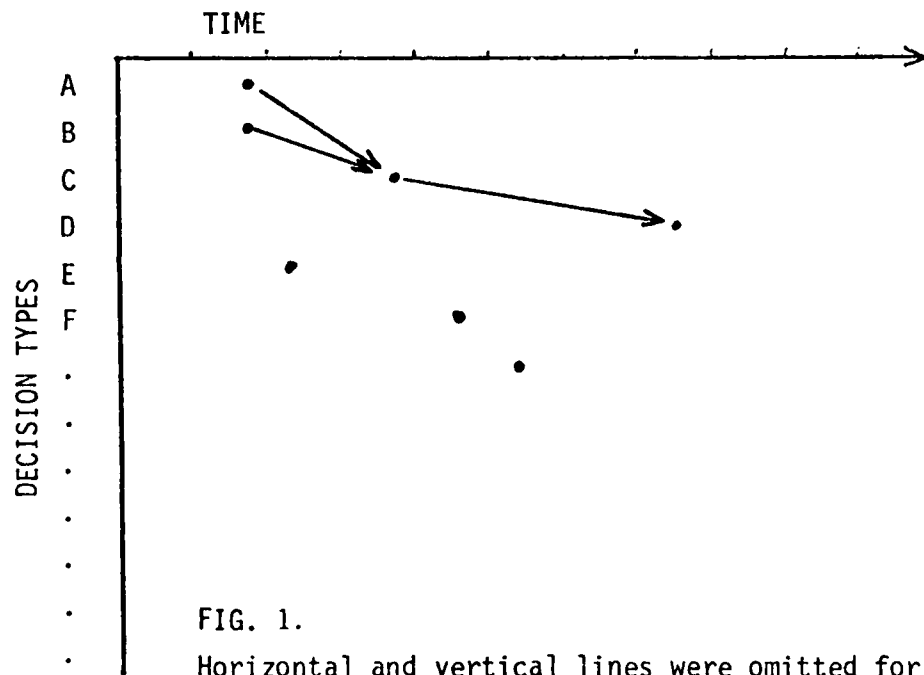


FIG. 1.

Horizontal and vertical lines were omitted for greater ease of communication.

Number, length and interconnectedness of forward diagonals will be of importance in several of the measures that are discussed below.

³Such diagonal connections in the matrix will later be referred to as "integrations".

Diagonals are sometimes drawn with arrowhead facing backwards. If, for example, a decision maker or a decision-making group engages in action A without considering a future action, but later finds that action A is now of use when a later action is decided upon, a backward arrow diagonal between the later action and the previous action may be drawn. As a rule, interconnectedness among backward diagonals does not occur with great frequency.

(6) End Effects.

Whether or not a diagonal is drawn depends, of course, on whether a planned later decision is indeed produced as a follow-up to the earlier decision. Where a decision task ends abruptly, the opportunity to carry out that later decision may not exist. Such an event could arbitrarily limit the number of diagonals produced by decision makers as it is reflected in the obtained measures (see below). Where decision making is measured in experimental settings, randomization of time periods reflecting or containing potentially differing environmental conditions may be of importance to avoid a constant error. Calculations of probabilities of diagonal connections may be utilized as well (see the measurement section below).

(7) Establishing Relationships in the Matrix.

For purposes of analysis, it is important to establish clear relationships (a) between inputs and subsequent output decisions and (b) among decisions which are causally or strategically related (as shown in diagonals). The only perfect representation of these relationships exists in the brain of the decision maker(s) at the moment the relevant decisions are made. Any measure of those relationships can, consequently, be subject to some error. Clearly, it is important to opt for the least amount of error in any experimental or observational setting. Certainly, the error levels would likely be much smaller in a well designed complex experimental simulation than in an ongoing free environment. In experimental simulation, records of planning can be obtained during the planning process. In real-world task environments, such less precise techniques as post-hoc interviews are required.

Ideally, the decision maker(s) should be asked immediately (upon making a decision) to indicate (a) any information received upon which the decision is based, and (b) any planned subsequent decisions that they might employ as a follow-up to a current decision. With some efforts, this can be achieved in complex experimental simulations (the participants may have to be persuaded, however, that indicating previously received information and indicating planned future decisions would be of value to them in terms of long range outcome). In many free simulations (particularly if interrupt control is lacking) and in the observation of real-world decision-making environments these kinds of questions cannot be asked. Collecting data from participants in the decision-making task after completion of the task often introduces serious bias. The only remaining solution requires experts to consider the decisions made and to judge whether these decisions were responses to previous information and/or were part of a decision-making sequence that should be represented by diagonal connections. Hopefully, interjudge reliability for such a task is high. Previous experience has shown that judges produce little variable error in making their judgements. As long as the judges have no particular biases for or against certain decision makers which they are evaluating, constant errors across the various samples result in little comparison error among decision makers or decision-making groups.

Establishing connections between inputs and decisions on the basis of expert judgments is relatively easy. Respondent decisions are typically directly related to the verbal content of the information received or describe the same location or information source contained in the input. When such commonalities are seen, a connection may be assumed to exist. More difficult is the establishment of connections among decisions. Obviously, where one decision directly refers to a previous decision ("Order the unit which we moved to quadrant X5 to fire on.....") a diagonal connection is directly established. However, is this a forward or a backward diagonal? If we were able to ask the decisions maker(s) about future decisions, when the decision to move the unit was made, then we do now know. If we were not able to ask (in free simulations of the

type described above or in real-world applied decision-making settings), then we cannot know. In these cases forward and backward diagonals cannot be distinguished and arrowheads cannot be drawn.

Let us return to the decision-making settings where relationships among decisions (connections) must be judged by expert observers of the decision-making sequence. Where no clear relationship is stated by the decision maker(s), aids must be used to determine whether relationships do likely exist. Such commonalities among decisions as addressee, location, action etc. are useful for this purpose. The most reliable of these is probably location. In a military setting, to give a relatively simple example, moving artillery to quadrant X5, asking it to fire on Y4, moving infantry to Y5 and finally ordering the infantry to attempt to take X5 would reflect a series of interrelated decisions across time. It should be noted here that moving troops to Y5 and another troop unit (both infantry) to X5 (at a later time) would not result in a diagonal connection: both represent the same decision type. This outcome is intentional: repetitious action is not necessarily representative of strategic action. If, on the other hand, both units are later asked to attack Y4, connections between the two movements and the later attack would be drawn. Obviously, a few decision sequences may be difficult to judge in terms of their potential interconnectedness. To the degree to which the judge can develop a picture of the strategy decision makers used (or if the judge can obtain advance information about their plans), the determination of strategic relationships will be considerably easier. In any case, if, after considerable thought, the judge is uncertain whether two decisions are or are not related to each other, it is better to err by omission. Uncertain relationships (interconnections) should not be scored to avoid artificially inflating some of the measures (below) which can show near quadratic effects of erroneously scored relationships.

An example of two decision matrices is provided on page 10. The figure shows decision matrices by two groups which differed in their decision-making styles (complexity).

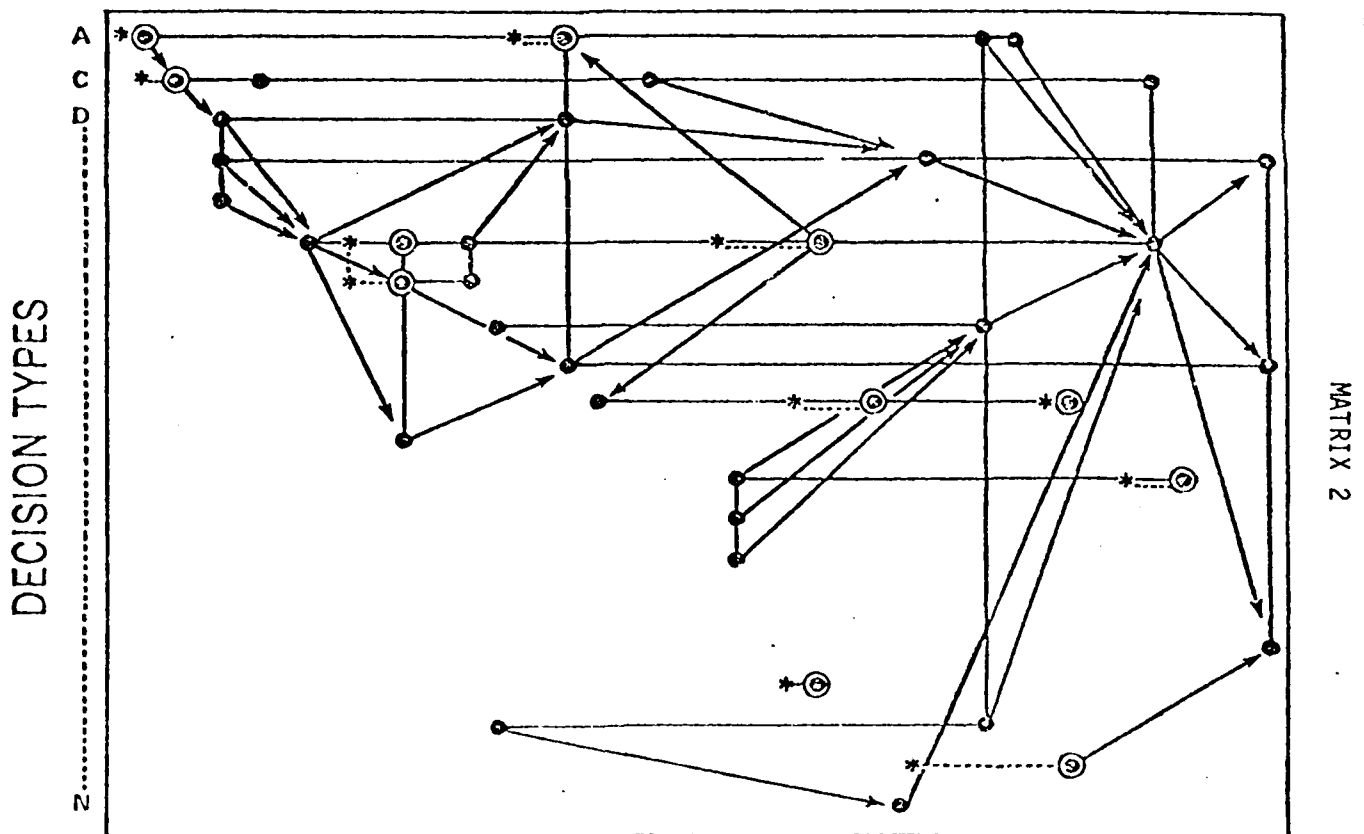
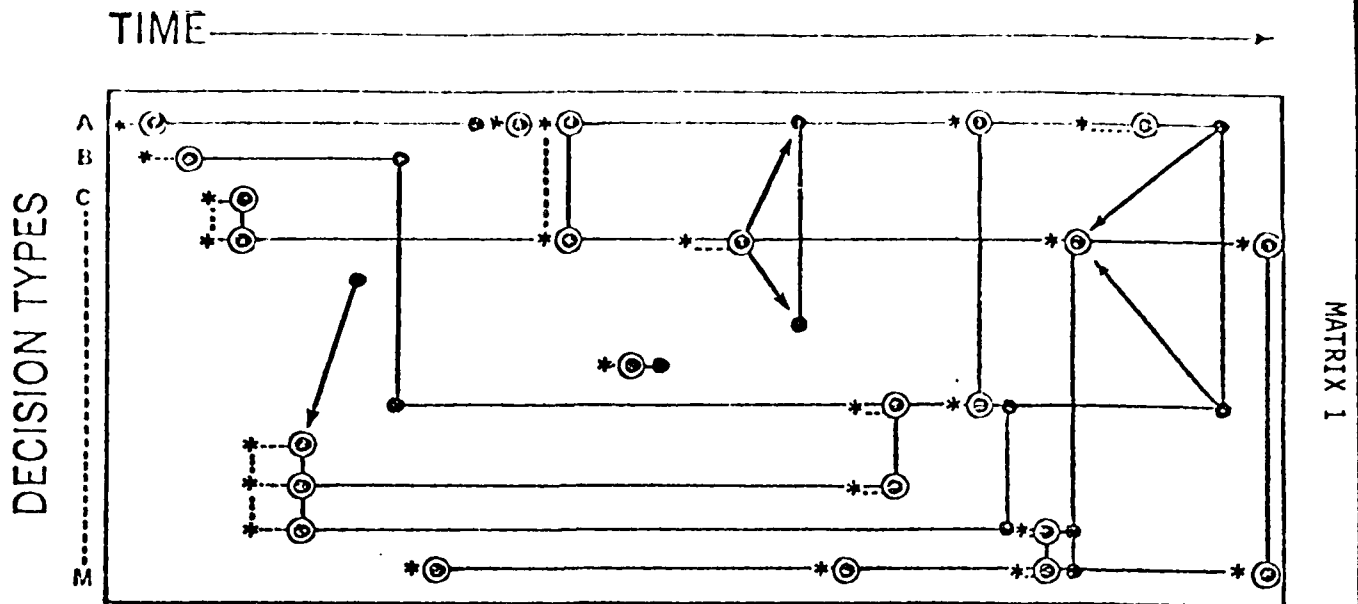


FIG. 2.

EACH POINT REPRESENTS A DECISION
 EACH VERTICAL LINE CONNECTS DECISIONS MADE AT THE SAME POINT IN TIME.
 EACH HORIZONTAL LINE CONNECTS DECISIONS OF THE SAME TYPE MADE AT DIFFERENT POINTS IN TIME.
 EACH DIAGONAL REPRESENTS THE STRATEGIC INTEGRATION OF DIFFERENT DECISIONS AT DIFFERENT POINTS IN TIME. DIAGONALS POINTING FORWARD REFLECT ADVANCE STRATEGIC PLANNING.
 EACH CIRCLED DOT REPRESENTS A DECISION RESPONSE TO INFORMATION RECEIVED AT * THE DOTTED DISTANCE FROM * TO ○ REFLECTS THE INFORMATION TO DECISION INTERVAL.
 EACH DECISION TYPE REPRESENTS A SELF SELECTED DIFFERENTIATED DECISION CATEGORY BASED ON AVAILABLE RESOURCES

MEASURES

As stated before, the measures to be presented here are generally orthogonal to each other and have been shown to be reliable and valid (criterion validity) for a number of task settings. Additional measures can be developed if they are useful for a specific task at hand. Calculation of the measures assumes that the time/event matrix has been drawn (by computer, if obtained from an experimental simulation, or by judges, if obtained from a free simulation or a real-world decision-making environment). The various measures reflect different kinds of task performance. In-and-of-themselves, each measure cannot be considered a reflection of "good" vs. "bad" performance with regard to any particular criterion without considering momentary demands (e.g. environmental conditions). Without question, there are situations where long range planning (as reflected in the QIS measure, below) is of considerable value, and there are situations where such planning would be superfluous and inappropriate since task demands may require immediate (e.g. respondent, see below) actions. Each measure and its purpose is discussed below. Where necessary, examples of how to calculate these measures will be provided. For convenience of communication, we will again use decision making as our example. It should be remembered, however, that many other effects may be measured instead (or in addition, in 3-or n-dimensional matrices).

1. Number of Decision Categories. (Measure 3)⁴

This measure is a simple count of the number of decision categories which decision makers use during any specified time period. Any category which is part of the count may have been used once or more than once. The measure reflects whether a decision maker is likely to select smaller or larger numbers of action types. In addition, further analysis could reveal whether decision maker(s) are likely to select certain specific actions

⁴For the convenience of those who wish to use the micro-computer based simulations developed by Streufert, Swezey and associates, the measure number printed by the scoring program is provided in parentheses.

and eliminate others from consideration. The basic measure may be written as

$$\sum_{1}^P C$$

where C is the number of categories employed and 1 through P is the time period of participation in the simulation that is of interest for analysis and interpretation.

2. Spread across Decision Categories.⁵ (Measure 15)

$$\sum_{1}^P 2(d_{Ca} - d_{Cb}) + (d_{Cd} - d_{Ce})$$

where, d is the number of decisions

d_{Ca} is the number of decisions from the category or categories representing the upper ten percent of decision frequency,

d_{Cb} is the number of decisions from the category or categories representing the lowest ten percent of decision frequency,

d_{Cd} is the number of decisions from the category or categories representing the upper fifty percent of decision frequency, and

d_{Ce} is the number of decisions from the category or categories representing the lower fifty percent of decision frequency.

A high value of this measure suggests that the decision maker(s), while not necessarily totally ignoring potential decision categories, is nonetheless spending the major effort on a limited number of activities.

⁵This measure has been slightly modified compared to previous applications.

For example, an executive who makes most of the decisions specifically related to profit or a military commander who lets the infantry do nearly all of the fighting without considerable support by other units would score high on this measure. A low score, on the other hand, would reflect a more well rounded approach to a decision-making problem. The measure is not meaningful if only one decision category is utilized.

Example.

Let us assume that decision maker(s) made a total of fifty decisions during a given period of time. These decisions represented the following decision types:

Decision Type	Number of decisions
A	5
B	10
C	20
D	3
E	8
F	1
G	<u>3</u>
Total:	50

The upper and lower ten percent would represent a value of 5 each (10% of fifty) while the upper and lower 50% would represent values of 25 each.

Decision Type Category C is included in the upper 10%, providing a value of 20 (the number of decisions in that decision type category) for d_{Ca} .

Decision Type Categories D, F and G are included in the lower ten percent providing a value of 7 for d_{Cb} .

Decision Categories B and C are included in the upper 50%, providing a value of 30 for d_{Cd} .

Decision Categories B, E, A, D, F and G are included in the lower 50% providing a value of 30 for d_{Ce} .

The resulting calculation for spread across decision categories is then:

$$2 (20 - 7) + (30 - 30) = 26.$$

Another, somewhat simpler way of calculating spread involves averaging the frequency counts of all decision categories utilized.

3. Average Spread Across Decision Categories. (Measure 16)

The formula for Spread Across Decision Categories is to some degree affected by the number of categories used (and/or available). To correct for potential errors, particularly when many decision categories are available and utilization differs widely among individuals or groups, an additional measure was introduced. This measure divides the score for spread across categories by the number of categories utilized. The average spread measure is not assumed (or demonstrated) to be orthogonal from Spread Across Decision Categories. It may be calculated as

$$\frac{\sum_{i=1}^P 2(d_{Ca} - d_{Cb}) + (d_{Cd} - d_{Ce})}{\sum_{i=1}^P C}$$

For the example of decision types and decision frequencies on page the obtained numerical score would be 26 (the value obtained for the Spread Across Decision Categories measure) divided by the Number of Categories measure, here 7 = 3.7143.

4. Number of Decisions. (Measure 1)

This measure reflects the amount of decision making activity. It consists of a count of the number of decisions made, i.e. the number of points in the matrix.

5. Number of Integrations. (Measures 4 and 7)

$$\sum_{l=1}^P i_f \quad \text{or} \quad \sum_{l=1}^P (i_b) \quad \text{or} \quad \sum_{l=1}^P (i_b + i_f) = \sum_{l=1}^P i$$

where i_f are forward integrations (relationships, i.e. connections among decision-making points with diagonal arrows pointing forward),

i_b are backward integrations (relationships, i.e. connections among decision-making points with diagonal arrows pointing backward),

i are integrations i.e. relationships where directionality cannot be established.

As discussed earlier, some decision-making tasks (particularly real-world decision-making settings where the researcher or observer cannot interfere) do not lend themselves to questioning the decision maker(s) about the intent of decisions. Consequently it may be impossible to determine whether a connection (relationship) among decisions reflects forward integrations (planning a later decision at the time an earlier decision has been made), or reflects backward integration, (using a previous decision to advantage although the connection was not considered at the time the earlier decision did occur). However, where possible, forward diagonals in the matrix should be counted as forward integrations; backward diagonals should be counted as backward integrations. Translation of diagonals into integration scores is achieved on a one-to-one basis: Counting the number of diagonals of a specific type produces the relevant integration score. Where no distinction between forward and backward diagonals could be made, integrations are counted without concern for the direction of arrowheads.

Example.

For simplicity's sake, let us return to the example matrix in Figure 2, page 10. The upper matrix contains two forward diagonals, i.e. a score of 2 for i_f (forward integrations). It contains three backward diagonals,

i.e. a score of 3 for i_b (backward integrations). The score for i ($i_f + i_b$) would be 5. Obviously the score for the lower matrix in figure 2 is considerably higher.

The measures concerned with number of integrations, integration time weight and QIS depend on the diagonals connecting an earlier and a subsequent decision, indicating that the earlier decision made the later decision possible (strategic time sequence). Where an entire matrix is analyzed, simple counting or statistical processing of the number of diagonals is sufficient. However, if an experimenter or observer is concerned with a limited time period as part of a larger decision time sequence (e.g. if different experimental conditions are introduced into an experimental simulation or if artificial or natural probes are utilized in a free simulation), diagonals will often cross the time lines that describe a period of interest. In that case, diagonals are counted as part of the time period during which they originate. If distinctions between backward and forward integrations (diagonals with backward and forward arrows) can be made, then backward integrations will be credited to the period of the subsequent (of two) decisions. Forward integrations will be credited to the period of the initial (of two interconnected) decisions. If no distinctions between forward and backward diagonals will be made, all diagonals are credited as integrations to the initial decision. Again, enough time should be permitted after the last probe to allow integrations (connected to earlier decisions) to actually occur. Similarly, if specific experimental variations are utilized in experimental simulations, the variations (if manipulated "within" across time) should be randomized to replace constant error with variable error in the research design. Adjustments to reduce error due to the end of a measurement period are discussed below.

6. Integration Time Weight. (Measure 6)

$$\sum_{1}^P W$$

Where the measure for number of integrations was merely concerned with the frequency with which connections (strategic relationships) do occur among decisions, the measure of time weight considers the length of time of future planning. The measure utilizes the individual integrations discussed in the integration measure (diagonals) but measures each diagonal on the time dimension (in units chosen by the experimenter or observer, see above) and replaces the value of 1 (for the occurrence of the diagonal) with the value of the time length. Consider, for example, the matrix example in Figure 3, below:

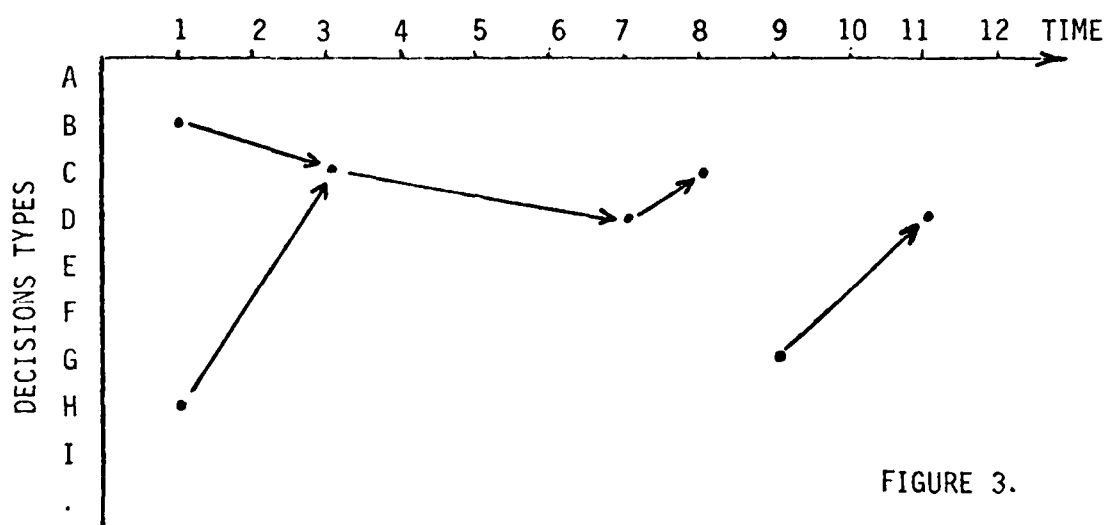


FIGURE 3.

The time weight for the forward integrations (diagonal connections) between the initial decisions B and H which are connected to decision C represents two time units each (remember that time units are selected by the experimenter/observer but must be held constant if numerical comparisons among scores for different decision makers are to be made). The connection between decisions C and D represents four time units. The connection between D and C represents one time unit, and finally the connection between G and D represents two time units. The total score for Integration time weight in this matrix is then $2 + 2 + 4 + 1 + 2 = 11$.

7. QIS Quality of Integrated Strategies. (Measure 9)

$$\sum_{i=1}^P W(1 + n_p + n_f)$$

Where W represents the length of the time dimension for any forward integration (or any integration, if distinctions between forward and backward integrations cannot be made). W is the last measure discussed above (integration time weight)

n_p is the number of other forward integrations (or any integration, if distinctions between forward and backward integrations cannot be made) connecting to the decision point representing the initial decision in a diagonal connection between two decisions and

n_f is the number of forward integrations (or any integrations, if distinctions between forward and backward integrations cannot be made) connecting to the decision point representing the subsequent decision in a diagonal connection between two decisions.

The number of integrations n_p and n_f here include only those integrations which are directly connected to either the initial (n_p) or subsequent (n_f) decision points.

The QIS measure is concerned with the degree to which planning (strategic behavior) follows an overall pattern vs. is composed of a number of separate unrelated plans. While the score for number of integrations may, for example, be the same in either case, an overall plan connecting all components of the decision-making sequence in a combined strategy would result in a higher; separate strategic plans would result in lower QIS scores. QIS measures tend to distinguish between decision-making quality more when decision makers operate at advanced decision-making levels. QIS scores would not exceed Integration Time Weight scores when all integrations are made without reference to each other, i.e. where an overall strategic plan does not exist or is not developed.

Example.

Let us again return to Figure 3, page 17. A QIS value would be established for each diagonal in the matrix. Let us initially take the diagonal which is connecting B and C. We already concluded that its

weight (W) score is 2. There are no diagonals connecting to its beginning point. On the other hand, there are two diagonals connecting directly to its end point. The score would be

$$2(1 + 0 + 2) = 6.$$

The same value of 6 would also be obtained for the H to C diagonal.

The C to D diagonal with a W value of 4 connects to two other diagonals at its beginning point and one other diagonal at its end point. Its score would be

$$4(1 + 2 + 1) = 16.$$

In turn the D to C connection would be

$$1(1 + 1 + 0) = 2.$$

Finally, the G to D diagonal maintains its W value since there are no diagonals connected to either the initial nor the subsequent decisions:

$$2(1 + 0 + 0) = 2.$$

For this matrix the total QIS score then would be

$$6 + 6 + 16 + 2 + 2 = 32.$$

8. Multiplicity of Integration.

The multiplicity measure is closely related to the previous (QIS) measure. It does, however, not take time between the original decision and the planned future decision into account. While this measure is not considered to be orthogonal to the QIS measure, it is designed to be supplemental and potentially more meaningful than the QIS measure in situations where responding (including strategic - integrated - responding) must occur quite rapidly or where the time delay between original and subsequent decision is more a function of task demands than of decision planning characteristics. The formula for multiplicity can be directly derived from the formula for Quality of Integrated Strategies by removing the time weight term W:

$$\sum_{i=1}^P (1 + n_p + n_f)$$

9. Weighted QIS. (Measure 10)

Weighted Quality of Integrated Strategies (WQIS) is an extension of the QIS measure to obtain scores for the sequential chain of interconnections among integrated decisions over long periods of time (i.e. multiple long term strategic actions that are coordinated). Where the QIS formula calculated the time weight for an integration (diagonal connection between decisions points differing in time) and multiplied that weight value with the number of other diagonals connected directly to the beginning point (initial decision) and the end point (later integrated decisions) of an integration, the WQIS measure considers all integrations (diagonals) which lead in chain sequence to the decision which begins the integration, and all integrations (diagonals) which follow the later integrated decision, as long as there is no interruption in diagonal (integration) links. Because of the multiplicative nature of this measure, quite high scores can be obtained when additional links are added in any strategic chain of decisions. Where no more than three decision points (differing in time) are connected with diagonals (integrations), the WQIS measure will not differ from the QIS measure. Where four decision points (three sequential diagonals) are involved, the measure will not differ for the middle integration, but will differ for the outer two integration diagonals. With an even greater numbers of diagonal connections in chain sequence the score for WQIS would considerably exceed the QIS score. The formula for WQIS can be written as

$$\sum_{i=1}^P W(1 + n_{pp} + n_{ff})$$

where

n_{pp} is the number of forward integrations reflected in the term n_p for the QIS measure plus all other forward integrations connecting to these integrations, and so forth, until all integrations (diagonals in the matrix) which connect to each other and can be traced without interruption to the beginning point of the forward integration of interest have been exhausted.

n_{ff} is the number of forward integration reflected in the term n_f for the QIS measure plus all other forward integrations connecting to these integrations, and so forth, until all integrations (diagonals in the matrix) which connect to each other and can be traced without interruption to the later decision have been exhausted.

all other terms are the same used in previous formulas.

For the example in Figure 3 on page 17 the WQIS score would be $2(1+0+3) + 2(1+0+3) + 4(1+2+1) + 1(1+3+0) + 2(1+0+0) = 38$.

10. Number of Respondent Decisions. (Measure 2, without time limitation, see below)

$$\sum_{1}^P r$$

where r is any decision made within a given time period (see below) after receipt of relevant information, if that decision is made in direct response to the information.

Whether a decision is made in response to previously received information should ideally (as discussed earlier) be determined via an immediate response of the decision maker(s) but may have to be determined by competent judges if access to the decision maker(s) for questioning is not available.

Different decision-making situations require diverse time frames for the processing and accessing/communicating of information and subsequent decisions. Respondent decisions (as defined here, see some potential modifications below) are made quickly in response to incoming information. The decision output is usually not extensively pondered or considered in terms of some existing or emerging plan (strategy). For example, a respondent decision to the intrusion of enemy aircraft into friendly airspace may involve immediate defensive action. Certainly the reasons for that intrusion may be considered subsequently and may be reflected in future activities that may or may not be strategic (reflective of planning). Nonetheless the initial action, occurring as immediately as

possible, represents (often quite appropriate) one-to-one responding to the information. The time limitation between receipt of information and response which determines whether a decision was made quickly enough after receipt of that information to qualify as a respondent decision must depend on the constraints of the decision-making situation. In other words, that time frame must be determined individually for each group of decision-making settings of interest. Nonetheless that time frame cannot be changed from one decision-making measurement to the next, if comparisons are to be made.

As stated on page 5, information inputs are located in the time/event matrix in front of all relevant actions (on all relevant horizontals) as * or star points. Measurement of the length of time between input and decision follows the same time scale used for determining the weight of diagonals (page 14). Respondent decision values for each input-decision time sequence are added to a total value. Several decisions in response to a single item of information (if made within the set time limit) are counted separately.

It should be noted that two variants of the respondent decision-making measure have been used with success. One is "retaliatory decision making". In this measure sequences between information receipt and decision are not included in the score for "r" if the decision is connected with other decisions by a forward or backward diagonal. As modified, the measure of retaliatory decision making provides an estimate of non-strategic respondent behavior. Another modification of the respondent decision-making measure eliminates the time constraint on the information-decision sequence. Here all decisions made in response to information (no matter how much delayed) are counted. Such a measure determines the amount of respondent activity. It may be further modified by dividing that value by the total number of decisions to obtain the degree to which decision maker(s) behave in a more responsive way (as opposed to taking more of an initiative). Note, however, that the modifications of the respondent decision-making measure are not statistically independent (orthogonal) of each other. Nonetheless they can be quite useful for specific research or observation intents.

11. Average response speed. (Measure 11)

$$\frac{\sum_{l=1}^p t_r}{r_p}$$

where t_r is the elapsed time between information receipt and subsequent respondent decision, and

r_p is the number of respondent decisions made in the time period between 1 and p.

The response speed measure simply reflects the rapidity with which decision maker(s) respond to incoming information with respondent decisions. The time length between each input and the subsequent decision is measured; the sum of those measures is divided by the number of responses to information. For this measure it is worth while to consider a value for r (number of respondent decisions) which is not constrained by a time limitation between information receipt and subsequent decision.

12. Serial Connections. (Measure 12)

The serial connection measure is similar to the Number of Integrations measure but counts interconnections between decisions that are placed into the same decision category. For example, if decision makers decide to move troop unit A and plan to subsequently move troop unit B (and, when movement of B is carried out recall that they moved A as an antecedent to the movement of B) then a forward serial connection is established: Both decisions fall into a single decision category: troop movement. They are, by themselves, not likely to reflect an ongoing strategy unless they are also interconnected with other decisions drawn from different categories (to which they would be connected with diagonals in the matrix). Serial connections without integrations often reflect a stagnating series of moves that may fail to take complexities of the task environment into account. If associated with strategic moves (as reflected in high scores on such measures as Number of Integrations or QIS), they may be part of a general (e.g., in the military, an encircling) strategy.

Serial connections may be measured (as were number of integrations) in terms of forward, backward or general connections between decisions of a single category:

$$\sum_{1}^P i_{sf} \quad \text{or} \quad \sum_{1}^P i_{sb} \quad \text{or} \quad \sum_{1}^P (i_{sf} + i_{sb}) = \sum_{1}^P i_s$$

where i_{sf} are forward serial connections, and i_{sb} are backward serial connections.

13. Planned Integrations. (Measure 13)

Not all actions, here decisions, which are planned as a follow-up to current actions are actually carried out in the future. Time demands, changed situations, forgetfulness, new strategies and more may be the cause of lacking follow-up actions. In some cases, an incomplete connection between a current action and a planned future action may indicate poor strategy. The Planned Integrations measure reflects the number of times decision makers fail to carry out a subsequent future action that had been planned previously. The formula for planned integrations can be written as

$$\sum_{1}^P i_{pf}$$

where i_{pf} is a planned forward integration which was not carried out in the future.

Planned integrations that did not come to fruition may be compared with the number of integrations which were completed to obtain an estimate of the degree to which decision makers do operationalize their plans. This score would be reflected by the formula

$$\sum_{1}^P i_f - \sum_{1}^P i_{pf}$$

Finally, the planned integration measure may be utilized to estimate the assumed time value for number of integrations where that measure is truncated by the end of a measurement or observation sequence (e.g. at

final participation periods in experimental simulations or at the retirement of an executive or officer prior to final completion of a task). Under these conditions, it may not have been possible to complete all future decisions which were planned when a previous action was initiated. As a result the uncorrected measure for number of integrations would underestimate the strategic planning of the decision maker. This correction may be calculated as:

$$X = 1 - \frac{\sum_{l=1}^c i_{pf}}{\sum_{l=1}^c i_f}$$

where 1 through C is any prior time period (or periods) to which a time period under analysis is to be compared.

The obtained value of this correction is then multiplied with the total number of intended integrations:

$$X \left(\sum_{l=1}^P i_{pf} + \sum_{l=1}^P i_f \right)$$

to obtain the estimated value of corrected Number of Integrations. Unless the corrected value is less than the actually obtained score for Number of Integrations, the Number of Integrations score may be replaced by the corrected score. Similar calculations may be employed to correct other measures which are based on forward or backward integrations.

14. Multiplexity F. (Measure 5)

The Multiplexity F measure has some similarity to the Weighted QIS measure but differs from that measure in two ways: (1) similarly to the Multiplicity of Integration calculation it does not take the time between the original decision and the planned future decision into account, and (2) it focusses only on plans which are related to or are subsequent to the planned future decision. In other words, the measure is concerned with

the complexity of future strategies (viewed from any one present point in time) only. By necessity, this measure is truncated by the limitations imposed by time: where the participant in the task is forcibly removed from his setting within the near future from any one action or where the task is about complete, the Multiplexity F measure is necessarily going to produce a lower score. When the degree of task completion or when time to separation from the task is taken into consideration, such a measure can be quite valuable as it stands. However, where the measure is to stand representative for general performance at any point in a task, it is worth while to divide Multiplexity F by the time to completion of the task (possibly truncated by a maximum value which has to be developed on the basis of maximum possible planning in the particular task at hand.

Multiplexity F may be written as:

$$\sum_{l=1}^P (1 + n_{ff})$$

where n_{ff} is the number of forward integrations reflected in the term n_f of the QIS measure plus all other forward integrations connecting to these integrations which either connect to the planned future decision or originate in that future decision and later decisions to which the future decision is connected, until all integrations (diagonals in the matrix) which connect to the planned future decision or represent subsequent serially connected diagonals (without interruption) are exhausted.

To measure general multiplexity, the formula may be modified:

$$\frac{\sum_{l=1}^P (1 + n_{ff})}{t_r}$$

where t_r is the time remaining in the task.

15. Measures of Performance Quality

Criterion performance quality measures are not immediately obtained from the time/event matrix. While the matrix reflects the utilization of decision-making styles which form intermediate levels between general decision-making measures and criterion measures, they are not in-and-of-themselves measures of quality unless validated for specific decision-making settings. While validation has occurred for several research based and applied settings (e.g. executive decision making), it is strongly

suggested that the stylistic measures be re-validated for each setting and in each type of environmental stress condition. Previous research has shown that the measures described above respond considerably to (for example load) stressor conditions. As discussed earlier in this paper, different task requirements may also favor different measures (above) as correlates or representational values for performance quality criteria.

It should be remembered that we have been describing measures of task performance which are not based on the content quality of the actions taken. We have assumed, that the decision makers (or other organizational personnel) have been well trained in their respective tasks, and that outcome is going to be less likely based on knowledge of the task content, but more on how that knowledge is appropriately used. Where such assumptions cannot be made, additional measures of performance content will have to be introduced.

16. Additional measurement.

Again, as stated earlier in this paper, additional measures may be developed from the matrix as required by some research or observational intent for a specific task setting. Further, other measures may be related to or combined with (e.g. multiplicatively or as ratio values) with the measures utilized here. The time/event matrix is a generally useful tool to describe the task performance process (and sequence) of individuals and/or groups. There is, however, nothing sacred about the measures developed from the matrix. While they are independent of each other, such independence reflects statistical and mathematical rigor, not applied necessity. Varimax rotation to orthogonality (used for the development of the measures on the basis of raw data matrices) does not necessarily reflect all applied decision-making characteristics or settings. Measures obtained in future efforts may well be oblique to each other, or may interact with non-measured phenomena. Development of measures, based on the matrix, which best reflect the specific performance criterion of interest cannot be rejected on the basis of mathematical purity. The measures listed above have been shown to be quite robust in relationship to criterion performance. They will likely continue to be

robust; nonetheless modification of these techniques or development of new techniques based on the matrix is to be encouraged, where it is required be external criteria.

END

FILMED

7-83

DTIC